Jeffrey-Like Rules of Conditioning for the Dempster–Shafer Theory of Evidence

Hidetomo Ichihashi and Hideo Tanaka
University of Osaka Prefecture Osaka, Japan

ABSTRACT

Jeffrey's rule of conditioning is a rule for changing an additive probability distribution when the human perception of new evidence is obtained. It is a generalization of the normative Bayesian inference. Shafer showed how Jeffrey's generalization of Bayes' rule of conditioning can be reinterpreted in terms of the theory of belief functions. But Shafer's approach is different from the normative Bayesian approach and is not a straight generalization of Jeffrey's rule. There are situations in which we need inference rules that may well provide a convenient generalization of Jeffrey's rule. Therefore we propose new rules of conditioning motivated by the work of Dubois and Prade. Although the weak and strong conditioning rules of Dubois and Prade are generalizations of Bayesian conditioning, they fail to yield Jeffrey's rule as a special case. Jeffrey's rule is a direct consequence of a special case of our conditioning rules. Three kinds of normalizations in the rules of conditioning are discussed.

KEYWORDS: ***theory of evidence, Dempster's rule of combination, weak and strong conditioning, Jeffrey's rule of conditioning, upper and lower probabilities***

1. INTRODUCTION

For the purpose of devising reasoning techniques under uncertainty, people in artificial intelligence (Gordon and Shortliffe [1]) pay attention to the theory of evidence (Shafer [2]). Ishizuka et al. [3] applied this theory to the management of uncertainty in expert systems.

Address correspondence to H. Ichihashi, Department of Industrial Engineering, University of Osaka Prefecture, 4-804 Mozu-Umemachi, Sakai, Osaka 591, Japan.

International Journal of Approximate Reasoning 1989; 3:143–156

655 Avenue of the Americas, New York, NY 10010 0888-613X/89/$3.50

The unicity of Dempster's rule (Dempster [4]) for combining uncertain items of information issued from independent sources was proved by Dubois and Prade [5]. They also proposed weak and strong conditioning rules that are the generalization of conditioning [5]. Motivated by their work, this paper proposes three conditioning rules with normalization. Our rules are different from theirs in the way in which normalization is achieved.

Shafer [6] explained how Jeffrey's rule of conditioning can be understood in terms of belief functions (Shafer [2]). But Shafer's argument is based on the retrospective and constructive point of view. It is not a direct generalization of Jeffrey's rule. When the prior beliefs are additive and the new evidence bears only on a partition $E_1, E_2, \cdots, E_n$ of the frame Ω, then the new degrees of belief $P(E_i)$ obtained by Dempster's rule are different from those obtained by Jeffrey's rule. Neither the weak conditioning nor the strong conditioning rule of Dubois and Prade yields Jeffrey's rule as a special case. In our newly proposed rules of conditioning, Jeffrey's rule is a direct consequence of a special case. Our main concern is the normalization in the rule of conditioning.

2. BAYES' THEOREMS AND A GENERALIZATION BY JEFFREY

Let $\Theta = \{\theta_1, \theta_2, \cdots, \theta_k\}$ be a set of disjoint states of nature such as disease, and let $X = \{x_1, x_2, \cdots, x_l\}$ be a set of disjoint items of information such as information about symptoms. If the disease is θ_i, then the symptom is x_j with a probability $p(x_j | \theta_i)$. When a prior Bayesian belief function $p(\theta_i):\Theta \to [0, 1]$, which is to say, a prior probability function, is given, then a posterior Bayesian belief function assigns any particular Θ the degree of belief

$$p(\theta_i | x_j) = \frac{p(x_j | \theta_i) p(\theta_i)}{\sum_m p(x_j | \theta_m) p(\theta_m)} \tag{2.1}$$

Formula (2.1) is often called the Bayes theorem. Bayesian rules of inference for diagnosis are written for all $i \in \{1, \cdots, k\}$ and $j \in \{1, \cdots, l\}$ as follows.

- If the symptom is x_j, then the diagnosis is θ_i with posterior probability $p(\theta_i | x_j)$.

Further, if the symptom is given as a Bayesian belief function $p'(x):X \to [0, 1]$, then the posterior belief function is

$$p(\theta_i | p') = \sum_j p(\theta_i | x_j) p'(x_j) \tag{2.2}$$

for all i. Hence we have $\Sigma_i \, p(\theta_i | p') = 1$. The inference rules are written for all i, as follows.

- If the symptom is p', then the diagnosis is θ_i with a posterior probability $p(\theta_i | p')$.

Equation (2.2) is Jeffrey's rule of conditioning in its most general form. It should be noted that

$$p(x_j|p')=\sum_i p(\theta_i|x_j)p'(x_i)=p'(x_j) \tag{2.3}$$

More generally we have

$$p(E_j|p')=p'(E_j) \tag{2.4}$$

for the partition E_j of Ω where p' represents new evidence that bears directly only on the partition $E_1, E_2, \cdots, E_n$. The new probability of a set E_j is equal to the probability of E_j representing new evidence.

3. THE DEMPSTER–SHAFER THEORY OF EVIDENCE

Shafer's belief function (Shafer [2]) was originally called a lower probability by Dempster. A lower probability (Dempster [4]) is a mapping P_* from 2^Ω to [0, 1]. A lower probability is uniquely defined through the specification of basic probability assignment satisfying

$$p(\varnothing)=0, \qquad \sum_{B\subset\Omega} p(B)=1 \tag{3.1}$$

and we have

$$P_*(A)=\sum_{B\subset A} p(B), \qquad \forall A \in 2^\Omega \tag{3.2}$$

A set A such that $p(A) > 0$ is called a focal element. The upper probability $P^*(A) = 1 - P_*(\bar{A})$ is also defined as

$$P^*(A)=\sum_{B\cap A\neq\phi} p(B) \tag{3.3}$$

Suppose p_1 is the basic probability assignment for a lower probability P_{1*} over a frame Ω, and denote the focal elements of P_{1*} by $A_1, \cdots, A_k$. Also, the basic probability assignment of a second lower probability P_{2*} is p_2, and its focal elements are $B_1, \cdots, B_l$.

In order to carry out the combination of P_{1*} and P_{2*}, a probability mass of measure $p_1(A_i)\, p_2(B_j)$ is committed to the intersection of two sets A_i and B_j. The total probability mass exactly committed to a given subset A of Ω will have measure

$$\sum_{\substack{ij \\ A_i\cap B_j=A}} p_1(A_i)p_2(B_j)$$

The difficulty with this scheme is that it may happen that

$$\sum_{\substack{ij \\ A_i \cap B_j = \phi}} p_1(A_i)p_2(B_j) > 0 \tag{3.4}$$

A new basic probability assignment p_0 for the lower probability of P_{1*} and P_{2*} is defined by Dempster as

$$p_0(A) = \frac{\sum_{\substack{ij \\ A_i \cap B_j = A}} p_1(A_i)p_2(B_j)}{1 - \sum_{\substack{ij \\ A_i \cap B_j = \phi}} p_1(A_i)p_2(B_j)} \tag{3.5}$$

Let P_{1*} be a prior belief function, and let P_{2*} represent new evidence. Let Ω be the Cartesian product $\Theta \times X$. When p_1 is a regular probability assignment on $\Theta \times X$ and p_2 focuses on a single focal element $\{x_j\} \times \Theta$, then (3.5) is Bayes' rule of (2.1). Thus,

$$P_*(X \times \{\theta_i\}) = P^*(\mathrm{X} \times \{\theta_i\}) = p(\theta_i | x_j) \tag{3.6}$$

When p_2 is also a regular probability assignment on X that is a coarsening of $\Theta \times X$ (p_2 focuses on a partition of X), by (3.5) we have

$$p_0(\{x_j\} \times \{\theta_i\}) = \frac{p_2(\{x_j\} \times \Theta)p_1(\{x_j\} \times \{\theta_i\})}{1 - \sum_{\substack{i,j,m \\ \{x_m\} \times \Theta \cap \{x_j\} \times \{\theta_i\} = \phi}} p_2(\{x_m\} \times \Theta)p_1(\{x_j\} \times \{\theta_i\})} \tag{3.7}$$

Replacing $p_2(\{x_j\} \times \Theta)$ and $p_1(\{x_i\} \times \{\theta_j\})$ by $p'(x_j)$ and $p(x_j | \theta_i)\, p(\theta_i)$, respectively,

$$p_0(\{x_j\} \times \{\theta_i\}) = \frac{p'(x_j)p(x_j | \theta_i)p(\theta_i)}{1 - \sum_{\substack{i,j,m \\ \{x_m\} \times \Theta \cap \{x_j\} \times \{\theta_i\} = \phi}} p'(x_m)p(x_j | \theta_i)p(\theta_i)} \tag{3.8}$$

Hence

$$P_*(X \times \{\theta_i\}) = P^*(X \times \{\theta_i\}) = \sum_j p_0(\{x_j\} \times \{\theta_i\})$$
$$= \frac{\sum_j p(x_j | \theta_i)p(\theta_i)p'(x_j)}{K} \tag{3.9}$$

where

$$K = 1 - \sum_{\substack{i,j,m \\ \{x_m\} \times \Theta \cap \{x_j\} \times \{\theta_i\} = \phi}} p'(x_m)p(x_j | \theta_i)p(\theta_i)$$

Consequently we have

$$P_*(X\times\{\theta_i\})=P^*(X\times\{\theta_i\})\neq p(\theta_i|p') \tag{3.10}$$

where $p(\theta_i|p')$ is obtained by applying Jeffrey's rule. Thus (2.2) and (2.4) are not recovered by (3.5), or $P_*(E_i) \neq P_{2*}(E_i)$.

4. NEW RULES OF CONDITIONING

In order that Jeffrey's rule will be recovered directly when the prior belief is Bayesian and the new evidence bears directly only on the partition $E_1, E_2, \cdots, E_n$, we propose a new rule that gives a basic probability assignment such that the total probability mass will again have measure 1.

DEFINITION 1 *Let P_{1*} and P_{2*} be two lower probabilities over the same frame Ω, with basic probability assignments p_1 and p_2 and focal elements $A_1, \cdots, A_k$ and $B_1, \cdots, B_l$, respectively. Then the function $p_3: 2^\Omega \to [0, 1]$ is defined for all $A \subset \Omega$ as*

$$p_3(\varnothing)=0 \tag{4.1}$$

$$p_3(A)=\sum_{\substack{ij\\A_i\cap B_j=A}}\frac{p_1(A_i)p_2(B_j)}{1-\sum_{\substack{m\\A_m\cap B_j=\phi}} p_1(A_m)} \tag{4.2}$$

where it is assumed that for all $j \in \{1, \cdots, k\}$

$$\sum_{\substack{m\\A_m\cap B_j=\phi}} p_1(A_m)<1 \tag{4.3}$$

It is obvious that the function p_3 of (4.2) takes nonnegative values. We will show that the $p_3(A)$ sum to 1.

$$\sum_{A\subset\Omega} p_3(A)=p_3(\varnothing)+\sum_{\substack{A\subset\Omega\\A\neq\phi}} p_3(A)$$

$$=\sum_{\substack{A\subset\Omega\\A\neq\phi}}\sum_{\substack{ij\\A_i\cap B_j=A}}\frac{p_1(A_i)p_2(B_j)}{1-\sum_{\substack{m\\A_m\cap B_j=\phi}} p_1(A_m)}$$

$$=\sum_{\substack{A\subset\Omega\\A\neq\phi}}\sum_j\frac{\sum_{\substack{i\\A_i\cap B_j=A}} p_1(A_i)p_2(B_j)}{1-\sum_{\substack{m\\A_m\cap B_j=\phi}} p_1(A_m)}$$

$$= \sum_j \frac{\sum_{A_i \cap B_j \neq \phi} p_1(A_i)}{1 - \sum_{A_i \cap B_j = \phi} p_1(A_i)} p_2(B_j)$$

$$= \sum_j p_2(B_j) = 1 \tag{4.4}$$

Hence the function p_3 of Definition 1 is a basic probability assignment.

The lower probability given by p_3 is denoted $P_{1*} \oplus P_{2*}$. This rule of conditioning is no longer commutative, i.e., $P_{1*} \oplus P_{2*} \neq P_{2*} \oplus P_{1*}$. This property is not shared by Dempster's rule. The weak conditioning rule in Dubois and Prade [5] is symmetric in the mass function being combined and so is not equivalent to our asymmetric rule in Definition 1.

THEOREM 1 *Suppose P_{2*} is given as*

$$P_{2*}(A) = \begin{cases} 1 & \text{if } B \subset A \\ 0 & \text{if } B \not\subset A \end{cases} \tag{4.5}$$

for a particular subset $B \subset \Omega$, and P_{1*} is another lower probability over Ω. Then P_{2*} is combinable with P_{1*} if and only if $P_{1*}(B) < 1$. If P_{2*} is combinable with P_{1*}, then

$$P_{3*}(A) = \frac{P_{1*}(A \cup \bar{B}) - P_{1*}(\bar{B})}{1 - P_{1*}(\bar{B})} \tag{4.6}$$

and

$$P_3^*(A) = \frac{P_1^*(A \cap B)}{P_1^*(B)} \tag{4.7}$$

for all $A \subset \Omega$.

Proof Since B is the only focal element of P_{2*} and $p_2(B) = 1$, (4.2) yields

$$p_3(A) = \frac{\sum_{i,\, A_i \cap B = A} p_1(A_1)}{1 - P_{1*}(\bar{B})} \tag{4.8}$$

$$P_{3*}(A) = \sum_{D,\, D \subset A} p_3(D)$$

$$= \frac{\sum_{D,\, \phi = D \subset A} \sum_{i,\, D = A_i \cap B} p_1(A_i)}{1 - P_{1*}(\bar{B})}$$

$$= \frac{\sum_{\substack{A_i \subset A \cup \bar{B} \\ A_i \subset \bar{B}}} P_1(A_i)}{1 - p_{1*}(\bar{B})}$$

$$= \frac{P_{1*}(A \cup \bar{B}) - P_{1*}(\bar{B})}{1 - P_{1*}(\bar{B})} \tag{4.9}$$

Hence

$$\begin{aligned} P_3^*(A) &= 1 - P_{3*}(\bar{A}) \\ &= \frac{P_1^*(B) - P_{1*}(\bar{A} \cup \bar{B}) + P_{1*}(\bar{B})}{P_1^*(B)} \\ &= \frac{P_1^*(A \cap B)}{P_1^*(B)} \qquad \blacksquare \end{aligned} \tag{4.10}$$

We can call (4.7) the rule of conditioning by the upper probability P_1^*. This property is shared by Dempster's rule.

THEOREM 2 *When P_{1*} is given as*

$$P_{1*}(A) = \begin{cases} 1 & \text{if } B \subset A \\ 0 & \text{if } B \not\subset A \end{cases} \tag{4.11}$$

for a particular subset $B \subset \Omega$, and P_{2*} is another lower probability over Ω, then P_{2*} is combinable with P_{1*} if and only if for all focal elements B_j of P_{2*}, $B_j \cap B \neq \phi$, and (4.3) is reduced to

$$\sum_{A_m \cap B_j = \phi} p_1(A_m) = 0 \tag{4.12}$$

where $A_m = B$.

If P_{2*} is combinable with P_{1*}, then

$$P_{3*}(A) = P_{2*}(A \cup \bar{B}) \tag{4.13}$$

and

$$P_3^*(A) = P_2^*(A \cap B) \tag{4.14}$$

for all $A \subset \Omega$.

Proof Since (4.12) is assumed for all j,

$$\begin{aligned} p_3(A) &= \sum_{B \cap B_j = A} p_1(B) p_2(B_j) \\ &= \sum_{B \cap C = A} p_2(C) \end{aligned} \tag{4.15}$$

$$P_{3*}(A)=\sum_{D\subset A} p_3(D)=\sum_{\phi\neq D\subset A}\sum_{B\cap C=D} p_2(C)$$

$$=\sum_{\phi\neq B\cap C\subset A} p_2(C)=\sum_{\substack{C\subset A\cup\bar{B}\\ C\not\subset\bar{B}}} p_2(C)$$

$$=\sum_{C\subset A\cup\bar{B}} p_2(C)-\sum_{C\subset\bar{B}} p_2(C)=P_{2*}(A\cup\bar{B}) \tag{4.16}$$

Hence

$$P_3^*(A)=1-P_{3*}(\bar{A})=1-P_{2*}(\bar{A}\cup\bar{B})$$

$$=1-P_{2*}(\overline{A\cap B})=P_2^*(A\cap B) \quad \blacksquare \tag{4.17}$$

If P_{1*} and P_{2*} are given as in Theorem 1, and

$$\sum_{A_m\cap B=\phi} p_1(A_m)=0 \tag{4.18}$$

is assumed, then, by Theorem 1,

$$P_{3*}(A)=P_{1*}(A\cup\bar{B}) \tag{4.19}$$

and

$$P_3^*(A)=P_1^*(A\cap B) \tag{4.20}$$

Therefore in this case $P_{1*}\oplus P_{2*}=P_{2*}\oplus P_{1*}$ holds. In other words, the rule of combination is commutative.

Suppose Ω is the Cartesian product $\Theta\times X$, and p_1 and p_2 are the regular probability assignments on $\Theta\times X$ and X, respectively. Then we have

$$P_3(\{x_j\}\times\{\theta_i\})=\frac{p_2(\{x_j\}\times\Theta)\cdot p_1(\{x_j\}\times\{\theta_i\})}{1-\sum_{n\neq j,m} p_1(\{x_n\}\times\{\theta_m\})}$$

$$=\frac{p_2(\{x_j\}\times\Theta)\cdot p_1(\{x_j\}\times\{\theta_i\})}{\sum_m p_1(\{x_j\}\times\{\theta_m\})} \tag{4.21}$$

Replacing $p_2(\{x_j\}\times\Theta)$ and $p_1(\{x_j\}\times\{\theta_i\})$ by $p'(x_j)$ and $p(x_j|\theta_i)p(\theta_i)$, respectively, as in (3.7), we have

$$P_3(\{x_j\}\times\{\theta_i\})=\frac{p'(x_j)p(x_j|\theta_i)p(\theta_i)}{\sum_m p(x_j|\theta_m)p(\theta_m)}=p(\theta_i|x_j)p'(x_j) \tag{4.22}$$

Hence

$$P_3^*(X\times\{\theta_i\})=P_{3*}(X\times\{\theta_i\})=\sum_n p(\theta_i|x_n)p'(x_n)$$

$$=p(\theta_i|p') \tag{4.23}$$

and (2.2) and (2.4) are recovered by (4.2). Generally, $P_{3*}(E_i) = P_{2*}(E_i)$ holds. Neither Dempster's rule nor the weak conditioning rule yields Jeffrey's rule as a special case.

The conditioning rule of formula (4.2) gives higher priority to the second evidence than to the first evidence.

THEOREM 3 *When p_1 and p_2 are the regular probability assignments for the lower probabilities P_{1*} and P_{2*} over a frame Ω and P_{2*} is combinable with P_{1*}, then*

$$P_{1*} \oplus P_{2*}=P_{2*} \tag{4.24}$$

Proof Let $A_1, \cdots, A_k$ denote the elements (points) of Ω. Since P_{2*} is combinable with P_{1*} in (4.24), $p_2(A_i) > 0$ implies $p_1(A_i) > 0$, and we have

$$p_3(A_i)=\frac{p_1(A_i)p_2(A_i)}{1-\sum_{m,\ A_m\cap A_i=\phi} p_1(A_m)}=p_2(A_i) \tag{4.25}$$

for all focal elements A_i of P_{2*}. ■

This is the extreme case where the partition $E_1, E_2, \cdots, E_n$ is as fine as Ω itself (i.e., each E_i is a single point). This property is shared by the additive probability distribution in applying Jeffrey's rule of conditioning (see Shafer [6], p. 4).

DEFINITION 2 *Assuming that P_{1*} and P_{2*} are two lower probabilities as in Definition 1, the function $p_4:2^\Omega \to [0, 1]$ is defined as*

$$p_4(\varnothing)=0 \tag{4.26}$$

$$p_4(A)=\sum_{ij,\ A=A_i\subset B_j}\frac{p_1(A_i)p_2(B_j)}{1-\sum_{m,\ A_m\not\subset B_j} p_1(A_m)} \qquad \text{for all } A\subset\Omega \tag{4.27}$$

where

$$\sum_{m,\ A_m\not\subset B_j} p_1(A_m)<1 \qquad \text{for all } j\in\{1, \cdots, k\} \tag{4.28}$$

It is easy to prove that the function p_4 is a basic probability assignment. Let us now consider the rule of conditioning by p_4 in Definition 2.

THEOREM 4 *Let P_{1*} and P_{2*} be as in Theorem 1. Then P_{2*} is combinable with P_{1*} if and only if $P_1^*(\bar{B}) < 1$. If P_{2*} is combinable with P_{1*}, then*

$$P_{4*}(A)=\frac{P_{1*}(A \cap B)}{P_{1*}(B)} \tag{4.29}$$

and

$$P_4^*(A)=\frac{P_1^*(A \cup \bar{B})-P_1^*(\bar{B})}{1-P_1^*(\bar{B})} \tag{4.30}$$

for all $A \subset \Omega$.

Proof

$$p_4(A)=\sum_{A=A_i\subset B} {}_i \frac{p_1(A_i)}{1-\sum_{A_m\not\subset B} p_1(A_m)}$$

$$=\sum_{A=C\subset B} {}_C \frac{p_1(C)}{1-P_1^*(\bar{B})}$$

$$=\frac{\sum_{A=C\subset B} {}_C\, p_1(C)}{P_{1*}(B)} \tag{4.31}$$

$$P_{4*}(A)=\sum_{D\subset A} {}_D\, p_4(D)=\frac{\sum_{\phi\neq D\subset A} {}_D \sum_{D=C\subset B} {}_C\, p_1(C)}{P_*(B)}$$

$$=\frac{P_{1*}(A \cap B)}{P_{1*}(B)} \tag{4.32}$$

Hence

$$P_4^*(A)=1-P_{4*}(\bar{A})$$

$$=1-\frac{1-P_1^*(A \cup \bar{B})}{1-P_1^*(\bar{B})}$$

$$=\frac{P_1^*(A \cup \bar{B})-P_1^*(\bar{B})}{1-P_1^*(\bar{B})} \quad \blacksquare \tag{4.33}$$

Equation (4.29) is the rule of conditioning by the lower probability P_{1*}. Equation (4.29) is called the geometrical rule of conditioning.

When P_{1*} and P_{2*} are as in Theorem 2, then P_{2*} is combinable with P_{1*} if and only if for all focal elements $B_j \subset \Omega$ of P_{2*}, $B_j \subset B$. Hence P_{4*} focuses only on B. Thus $P_{4*}(A) = P_{1*}(A)$ and $P_4^*(A) = P_1^*(A)$.

When p_1 and p_2 are the regular probability assignments for P_{1*} and P_{2*}, and P_{2*} is combinable with P_{1*}, then we can readily see that $P_{1*} \oplus P_{2*} = P_{2*}$. As in (4.21) and (4.22), (2.2) is recovered by (4.27).

Definition 3 *Assuming that P_{1*} and P_{2*} are two lower probabilities as in Definition 1, the function p_5 is defined for all $A \subset \Omega$ as*

$$p_5(\varnothing)=0 \tag{4.34}$$

$$p_5(A)= \sum_{\substack{ij\\ A=A_i\\ A_i\cap B_j\neq\phi}} \frac{p_1(A_i)p_2(B_j)}{1-\sum_{\substack{m\\ A_m\cap B_j=\phi}} p_1(A_m)} \tag{4.35}$$

where it is assumed that for all j

$$\sum_{\substack{m\\ A_m\cap B_j=\phi}} p_1(A_m)<1 \tag{4.36}$$

p_5 is a basic probability assignment.

Theorem 5 *Let P_{1*} and P_{2*} be as in Theorem 1. Let $P_1^*(A, B)$ denote*

$$\sum_{\substack{C\\ A\cap C\neq\phi\\ B\cap C\neq\phi}} p_1(C)$$

If P_{2*} is combinable with P_{1*}, then

$$P_5^*(A)=\frac{P_1^*(A, B)}{P_1^*(B)} = \frac{P_1^*(A)+P_1^*(B)-P_1^*(A \cup B)}{P_1^*(B)} \tag{4.37}$$

and

$$P_{5*}(A)=\frac{P_{1*}(A)-P_{1*}(A \cap \bar{B})}{1-P_{1*}(\bar{B})} \tag{4.38}$$

for all A.

Proof

$$p_5(A)= \sum_{\substack{i\\ A=A_i\\ A_i\cap B\neq\phi}} \frac{p_1(A_i)}{1-\sum_{\substack{m\\ A_m\cap B=\phi}} p_1(A_m)} = \frac{\sum_{\substack{i\\ A=A_i\\ A_i\cap B\neq\phi}} p_1(A_i)}{P_1^*(B)} \tag{4.39}$$

$$P_5^*(A)=\sum_{\substack{D\\D\cap A\neq\phi}} p_5(D)$$

$$=\frac{\displaystyle\sum_{\substack{D\\D\cap A\neq\phi}}\ \sum_{\substack{i\\D=A_i\\A_i\cap B\neq\phi}} p_1(A_i)}{P_1^*(B)}=\frac{\displaystyle\sum_{\substack{i\\A\cap A_i\neq\phi\\B\cap A_i\neq\phi}} p_1(A_i)}{P_1^*(B)}$$

$$=\frac{P_1^*(A,B)}{P_1^*(B)} \tag{4.40}$$

$$P_1^*(A,B)=1-\sum_{\substack{i\\A_i\subset\bar{A}}} p_1(A_i)-\sum_{\substack{i\\A_i\subset\bar{B}}} p_1(A_i)+\sum_{\substack{i\\A_i\subset\bar{A}\cap\bar{B}}} p_1(A_i)$$

$$=1-P_{1*}(\bar{A})-P_{1*}(\bar{B})+P_{1*}(\bar{A}\cap\bar{B})$$

$$=P_1^*(A)+P_1^*(B)-P_1^*(A\cup B) \tag{4.41}$$

$$P_{5*}(A)=1-P_5^*(\bar{A})=1-\frac{P_1^*(\bar{A},B)}{P_1^*(B)}$$

$$=\frac{P_1^*(\bar{A}\cup B)-P_1^*(\bar{A})}{1-P_{1*}(\bar{B})}$$

$$=\frac{P_{1*}(A)-P_{1*}(A\cap\bar{B})}{1-P_{1*}(\bar{B})} \quad ■ \tag{4.42}$$

It should be noted that $P_1^*(A, B) \neq P_1^*(A \cap B)$.

We have proposed three conditioning rules. When the focal subsets of P_{1*} are singletons and P_{2*} focuses only on the partition $E_1, E_2, \cdots, E_n$, Jeffrey's rule of conditioning discussed in Section 2 is recovered in each case. The denominators of the conditioning rules represent measures of the extent of the conflict. Since P_{1*} and P_{2*} do not commit probability to disjoint (or contradictory) subsets A_i and B_j, the denominator of (4.2) in Definition 1 measures the extent of conflict in the sense that $A_i \cap B_j = \phi$, and that of (4.27) in Definition 2 measures the extent of the conflict in the sense that $A_i \not\subset B_j$. That of (4.35) is the same as Definition 1, but P_{1*} and P_{2*} commit a probability to A_i, if it is not contradictory to B_j. Hence we call the three conditioning rules plausible, credible, and possible conditioning, respectively. We can choose one of the three rules depending on the situation.

The drawback to the proposed rules of conditioning in Definitions 1-3 is that the condition of combinability is very restrictive. To relax the condition of combinability, we propose the following renormalizations.

DEFINITION 4 *Assuming that P_{1*} and P_{2*} are the same as in Definition 1,* *the functions p_3', p_4', and p_5' are defined for all A as*

$$p_3'(\varnothing)=0 \tag{4.43}$$

$$p_3'(A)=\frac{\sum_{ij,\ A_i\cap B_j=A}\left[p_1(A_i)p_2(B_j)\Big/\left(1-\sum_{m,\ A_m\cap B_j=\phi}p_1(A_m)\right)\right]}{1-\sum_{j,\ ((\cup_m A_m)\cap B_j)=\phi}p_2(B_j)} \tag{4.44}$$

$$p_4'(\varnothing)=0 \tag{4.45}$$

$$p_4'(A)=\frac{\sum_{ij,\ A=A_i\subset B_j}\left[p_1(A_i)p_2(B_j)\Big/\left(1-\sum_{m,\ A_m\not\subset B_j}p_1(A_m)\right)\right]}{1-\sum_{j,\ (\cup_{A_m\subset B_j}A_m)=\phi}p_2(B_j)} \tag{4.46}$$

$$p_5'(\varnothing)=0 \tag{4.47}$$

$$p_5'(A)=\frac{\sum_{ij,\ A=A_i,\ A_i\cap B_j\neq\phi}\left[p_1(A_i)p_2(B_j)\Big/\left(1-\sum_{m,\ A_m\cap B_j=\phi}p_1(A_m)\right)\right]}{1-\sum_{j,\ ((\cup_m A_m)\cap B_j)=\phi}p_2(B_j)} \tag{4.48}$$

where it is assumed that the denominator of each function is not equal to zero.

5. CONCLUDING REMARKS

We have proposed three rules of conditioning that are direct generalizations of Jeffrey's rule of conditioning. There might be situations in which we want to treat evidence asymmetrically, and in such cases our rules can provide a convenient generalization of Jeffrey's rule.

Our conditioning rules will be applied to decision problems treating the value of information sources in the framework of the theory of evidence.

ACKNOWLEDGMENTS

We are grateful to an anonymous referee for his valuable comments.

References

1. Gordon, J., and Shortliffe, E. H., A method for managing evidential reasoning in a hierarchical hypothesis space, *AI.* **26**, 323–357, 1985.

2. Shafer, G., *A Mathematical Theory of Evidence*, Princeton U. P., Princeton, N.J., 1976.

3. Ishizuka, M., Fu, K. S., and Yao, J. T. P., Inference procedures under uncertainty for the problem-reduction method, *Inf. Sci.* **28**(3), 179–206, 1982.

4. Dempster, A. P., Upper and lower probabilities induced by a multivalued mapping, *Ann. Math. Stat.* **38**(2), 325–339, 1968.

5. Dubois, D., and Prade, H., On the unicity of Dempster's rule of combination, *Int. J. Intell. Syst.* 1(2), 133–142, 1986.

6. Shafer, G., Jeffrey's rule of conditioning, Tech. Rep. No. 131, Stanford University, Stanford, Calif., 1979.